



IFW16

## RAW SEQUENCE LISTING

DATE: 07/20/2006

PATENT APPLICATION: US/09/611,949

TIME: 09:19:45

Input Set : A:\5470-451.ST25.TXT

Output Set: N:\CRF4\07202006\I611949.raw

3 <110> APPLICANT: Margolis, David M.  
 4 Romerio, Fabio  
 5 Coull, Jason J.  
 7 <120> TITLE OF INVENTION: HIV TRANSCRIPTION REPRESSOR COMPLEX AND COMPOSITIONS AND  
 METHODS  
 8 BASED THEREON  
 10 <130> FILE REFERENCE: 5470-451  
 12 <140> CURRENT APPLICATION NUMBER: US 09/611,949  
 13 <141> CURRENT FILING DATE: 2000-07-06  
 15 <150> PRIOR APPLICATION NUMBER: PCT/US98/00574  
 16 <151> PRIOR FILING DATE: 1998-01-13  
 18 <150> PRIOR APPLICATION NUMBER: US 60/036,242  
 19 <151> PRIOR FILING DATE: 1997-01-23  
 21 <160> NUMBER OF SEQ ID NOS: 7  
 23 <170> SOFTWARE: PatentIn version 3.3  
 25 <210> SEQ ID NO: 1  
 26 <211> LENGTH: 200  
 27 <212> TYPE: DNA  
 28 <213> ORGANISM: Human immunodeficiency virus  
 30 <400> SEQUENCE: 1  
 31 caagggactt tccgctgggg actttccagg gaggcgtggc ctgggcgga ctggggagtg 60  
 33 gcgagccctc agatgctgca tataagcagc tgctttttgc ctgtactggg tctctctggt 120  
 35 tagaccagat ttgagcctgg gagctctctg gctaactagg gaaccactg cttaagcctc 180  
 37 aataaagctt gccttgagtg 200  
 40 <210> SEQ ID NO: 2  
 41 <211> LENGTH: 2353  
 42 <212> TYPE: DNA  
 43 <213> ORGANISM: Homo sapiens  
 46 <220> FEATURE:  
 47 <221> NAME/KEY: CDS  
 48 <222> LOCATION: (241)..(1482)  
 50 <400> SEQUENCE: 2  
 51 cgccgagacg agcagcggcc gagcgcgcgc gggcgcgggc gcaccgaggc gagggaggcg 60  
 53 gggaagcccc gccgcgcgcg cccgcgccgc cccttcccc gccgcgccgc ccctctcccc 120  
 55 ccgcccgcgc gccgccttcc tccctctgcc ttccttcccc acggccggcc gcctcctcgc 180  
 57 ccgcccgcgc gcagccgagg agccgaggcc gccgcggccg tggcgcgga gccctcagcc 240  
 59 atg gcc tcg ggc gac acc ctc tac atc gcc acg gac ggc tcg gag atg 288  
 60 Met Ala Ser Gly Asp Thr Leu Tyr Ile Ala Thr Asp Gly Ser Glu Met  
 61 1 5 10 15  
 63 ccg gcc gag atc gtg gag ctg cac gag atc gag gtg gag acc atc ccg 336  
 64 Pro Ala Glu Ile Val Glu Leu His Glu Ile Glu Val Glu Thr Ile Pro  
 65 20 25 30  
 67 gtg gag acc atc gag acc aca gtg gtg ggc gag gag gag gag gag gac 384  
 68 Val Glu Thr Ile Glu Thr Thr Val Val Gly Glu Glu Glu Glu Glu Asp

## RAW SEQUENCE LISTING

DATE: 07/20/2006

PATENT APPLICATION: US/09/611,949

TIME: 09:19:45

Input Set : A:\5470-451.ST25.TXT

Output Set: N:\CRF4\07202006\I611949.raw

69	35	40	45	
71	gac gac gac gag gac ggc ggc ggt ggc gac cac ggc ggc ggg ggc ggc	432		
72	Asp Asp Asp Glu Asp Gly Gly Gly Gly Asp His Gly Gly Gly Gly Gly			
73	50	55	60	
75	cac ggg cac gcc ggc cac cac cac cac cat cac cac cac cac cac	480		
76	His Gly His Ala Gly His His His His His His His His His His			
77	65	70	75	80
79	ccg ccc atg atc gct ctg cag ccg ctg gtc acc gac gac ccg acc cag	528		
80	Pro Pro Met Ile Ala Leu Gln Pro Leu Val Thr Asp Asp Pro Thr Gln			
81	85	90	95	
83	gtg cac cac cac cag gag gtg atc ctg gtg cag acg cgc gag gag gtg	576		
84	Val His His His Gln Glu Val Ile Leu Val Gln Thr Arg Glu Glu Val			
85	100	105	110	
87	gtg ggc ggc gac gac tcg gac ggg ctg cgc gcc gag gac ggc ttc gag	624		
88	Val Gly Gly Asp Asp Ser Asp Gly Leu Arg Ala Glu Asp Gly Phe Glu			
89	115	120	125	
91	gat cag att ctc atc ccg gtg ccc gcg ccg gcc ggc ggc gac gac gac	672		
92	Asp Gln Ile Leu Ile Pro Val Pro Ala Pro Ala Gly Gly Asp Asp Asp			
93	130	135	140	
95	tac att gaa caa acg ctg gtc acc gtg gcg gcg gcc ggc aag agc ggc	720		
96	Tyr Ile Glu Gln Thr Leu Val Thr Val Ala Ala Ala Gly Lys Ser Gly			
97	145	150	155	160
99	ggc ggc ggc tcg tcg tcg tcg gga ggc ggc cgc gtc aag aag ggc ggc	768		
100	Gly Gly Gly Ser Ser Ser Gly Gly Arg Val Lys Lys Gly Gly			
101	165	170	175	
103	ggc aag aag agc ggc aag aag agt tac ctc agc ggc ggg gcc ggc gcg	816		
104	Gly Lys Lys Ser Gly Lys Lys Ser Tyr Leu Ser Gly Gly Ala Gly Ala			
105	180	185	190	
107	gcg ggc ggg cgc ggc gcc gac ccg ggc aac aag aag tgg gag cag aag	864		
108	Ala Gly Gly Arg Gly Ala Asp Pro Gly Asn Lys Lys Trp Glu Gln Lys			
109	195	200	205	
111	cag gtg cag atc aag acc ctg gag ggc gag ttc tcg gtc acc atg tgg	912		
112	Gln Val Gln Ile Lys Thr Leu Glu Gly Glu Phe Ser Val Thr Met Trp			
113	210	215	220	
115	tcc tca gat gaa aaa aaa gat att gac cat gag aca gtg gtt gaa gaa	960		
116	Ser Ser Asp Glu Lys Lys Asp Ile Asp His Glu Thr Val Val Glu Glu			
117	225	230	235	240
119	cag atc att gga gag aac tca cct cct gat tat tca gaa tat atg aca	1008		
120	Gln Ile Ile Gly Glu Asn Ser Pro Pro Asp Tyr Ser Glu Tyr Met Thr			
121	245	250	255	
123	gga aag aaa ctt cct cct gga gga ata cct ggc att gac ctc tca gat	1056		
124	Gly Lys Lys Leu Pro Pro Gly Gly Ile Pro Gly Ile Asp Leu Ser Asp			
125	260	265	270	
127	ccc aaa caa ctg gca gaa ttt gct aga atg aag cca aga aaa att aaa	1104		
128	Pro Lys Gln Leu Ala Glu Phe Ala Arg Met Lys Pro Arg Lys Ile Lys			
129	275	280	285	
131	gaa gat gat gct cca aga aca ata gct tgc cct cat aaa ggc tgc aca	1152		
132	Glu Asp Asp Ala Pro Arg Thr Ile Ala Cys Pro His Lys Gly Cys Thr			
133	290	295	300	

## RAW SEQUENCE LISTING

DATE: 07/20/2006

PATENT APPLICATION: US/09/611,949

TIME: 09:19:45

Input Set : A:\5470-451.ST25.TXT

Output Set: N:\CRF4\07202006\I611949.raw

```

135 aag atg ttc agg gat aac tcg gcc atg aga aaa cat ctg cac acc cac      1200
136 Lys Met Phe Arg Asp Asn Ser Ala Met Arg Lys His Leu His Thr His
137 305                               310                               315                               320
139 ggt ccc aga gtc cac gtc tgt gca gaa tgt ggc aaa gct ttt gtt gag      1248
140 Gly Pro Arg Val His Val Cys Ala Glu Cys Gly Lys Ala Phe Val Glu
141                               325                               330                               335
143 agt tca aaa cta aaa cga cac caa ctg gtt cat act gga gag aag ccc      1296
144 Ser Ser Lys Leu Lys Arg His Gln Leu Val His Thr Gly Glu Lys Pro
145                               340                               345                               350
147 ttt cag tgc acg ttc gaa ggc tgt ggg aaa cgc ttt tca ctg gac ttc      1344
148 Phe Gln Cys Thr Phe Glu Gly Cys Gly Lys Arg Phe Ser Leu Asp Phe
149                               355                               360                               365
151 aat ttg cgc aca cat gtg cga atc cat acc gga gac agg ccc tat gtg      1392
152 Asn Leu Arg Thr His Val Arg Ile His Thr Gly Asp Arg Pro Tyr Val
153                               370                               375                               380
155 tgc ccc ttc gat ggt tgt aat aag aag ttt gct cag tca act aac ctg      1440
156 Cys Pro Phe Asp Gly Cys Asn Lys Lys Phe Ala Gln Ser Thr Asn Leu
157 385                               390                               395                               400
159 aaa tct cac atc tta aca cat gct aag gcc aaa aac aac cag      1482
160 Lys Ser His Ile Leu Thr His Ala Lys Ala Lys Asn Asn Gln
161                               405                               410
163 tgaaaagaag agagaagacc cttctcgacc acgggaagca tcttccagaa gtgtgattgg      1542
165 gaataaatat gcctctcctt tgtatattat ttctaggaag aattttaaaa atgaatccta      1602
167 cacacctaag ggacatgttt tgataaagta gtaaaaatta aaaaaaaaaa actttactaa      1662
169 gatgacattg ctaagatgct ctatcttgct ctgtaatctc gtttcaaaaa cacagtgttt      1722
171 ttgtaaagtg tgggtcccaac aggaggacaa ttcatgaact tcgcatcaaa agacaattct      1782
173 ttatacaaca gtgctaaaaa tgggacttct tttcacattc ttataaatat gaagctcacc      1842
175 tgttgcttac aattttttta attttgtatt ttccaagtgt gcatattgta cacttttttg      1902
177 gggatatgct tagtaatgct acgtgtgatt tttctggagg ttgataactt tgcttgagcgt      1962
179 agattttctt taaaagaatg ggcagttaca tgcatacttc aaaagtattt tcctgtaaaa      2022
181 aaaaaaaaaaag ttatataggt tttgtttgct atcttaattt tgggtgtatt ctttgatgtt      2082
183 aacacatttt gtataattgt atcgtagatc tgtattgaat catgtagtat caaatattag      2142
185 atgtgattta atagtgttaa tcaattttaa cccatttttag tcactttttt tttccaaaaa      2202
187 aatactgcca gatgctgatg ttcagtgtaa tttctttgcc tgttcagtta cagaaagtgg      2262
189 tgctcagttg tagaatgtat tgtacctttt aacacctgat gtgtacatcc catgtaacag      2322
191 aaagggaac aataaaatag caatcctaaa g                                2353
194 <210> SEQ ID NO: 3
195 <211> LENGTH: 414
196 <212> TYPE: PRT
197 <213> ORGANISM: Homo sapiens
199 <400> SEQUENCE: 3
201 Met Ala Ser Gly Asp Thr Leu Tyr Ile Ala Thr Asp Gly Ser Glu Met
202 1                               5                               10                               15
205 Pro Ala Glu Ile Val Glu Leu His Glu Ile Glu Val Glu Thr Ile Pro
206                               20                               25                               30
209 Val Glu Thr Ile Glu Thr Thr Val Val Gly Glu Glu Glu Glu Asp
210                               35                               40                               45
213 Asp Asp Asp Glu Asp Gly Gly Gly Gly Asp His Gly Gly Gly Gly Gly
214 50                               55                               60

```

## RAW SEQUENCE LISTING

DATE: 07/20/2006

PATENT APPLICATION: US/09/611,949

TIME: 09:19:45

Input Set : A:\5470-451.ST25.TXT

Output Set: N:\CRF4\07202006\I611949.raw

```

217 His Gly His Ala Gly His His His His His His His His His His
218 65              70              75              80
221 Pro Pro Met Ile Ala Leu Gln Pro Leu Val Thr Asp Asp Pro Thr Gln
222              85              90              95
225 Val His His His Gln Glu Val Ile Leu Val Gln Thr Arg Glu Glu Val
226              100             105             110
229 Val Gly Gly Asp Asp Ser Asp Gly Leu Arg Ala Glu Asp Gly Phe Glu
230              115             120             125
233 Asp Gln Ile Leu Ile Pro Val Pro Ala Pro Ala Gly Gly Asp Asp Asp
234              130             135             140
237 Tyr Ile Glu Gln Thr Leu Val Thr Val Ala Ala Ala Gly Lys Ser Gly
238 145              150             155             160
241 Gly Gly Gly Ser Ser Ser Ser Gly Gly Gly Arg Val Lys Lys Gly Gly
242              165             170             175
245 Gly Lys Lys Ser Gly Lys Lys Ser Tyr Leu Ser Gly Gly Ala Gly Ala
246              180             185             190
249 Ala Gly Gly Arg Gly Ala Asp Pro Gly Asn Lys Lys Trp Glu Gln Lys
250              195             200             205
253 Gln Val Gln Ile Lys Thr Leu Glu Gly Glu Phe Ser Val Thr Met Trp
254              210             215             220
257 Ser Ser Asp Glu Lys Lys Asp Ile Asp His Glu Thr Val Val Glu Glu
258 225              230             235             240
261 Gln Ile Ile Gly Glu Asn Ser Pro Pro Asp Tyr Ser Glu Tyr Met Thr
262              245             250             255
265 Gly Lys Lys Leu Pro Pro Gly Gly Ile Pro Gly Ile Asp Leu Ser Asp
266              260             265             270
269 Pro Lys Gln Leu Ala Glu Phe Ala Arg Met Lys Pro Arg Lys Ile Lys
270              275             280             285
273 Glu Asp Asp Ala Pro Arg Thr Ile Ala Cys Pro His Lys Gly Cys Thr
274              290             295             300
277 Lys Met Phe Arg Asp Asn Ser Ala Met Arg Lys His Leu His Thr His
278 305              310             315             320
281 Gly Pro Arg Val His Val Cys Ala Glu Cys Gly Lys Ala Phe Val Glu
282              325             330             335
285 Ser Ser Lys Leu Lys Arg His Gln Leu Val His Thr Gly Glu Lys Pro
286              340             345             350
289 Phe Gln Cys Thr Phe Glu Gly Cys Gly Lys Arg Phe Ser Leu Asp Phe
290              355             360             365
293 Asn Leu Arg Thr His Val Arg Ile His Thr Gly Asp Arg Pro Tyr Val
294              370             375             380
297 Cys Pro Phe Asp Gly Cys Asn Lys Lys Phe Ala Gln Ser Thr Asn Leu
298 385              390             395             400
301 Lys Ser His Ile Leu Thr His Ala Lys Ala Lys Asn Asn Gln
302              405             410
305 <210> SEQ ID NO: 4
306 <211> LENGTH: 2418
307 <212> TYPE: DNA
308 <213> ORGANISM: Homo sapiens
311 <220> FEATURE:

```

## RAW SEQUENCE LISTING

DATE: 07/20/2006

PATENT APPLICATION: US/09/611,949

TIME: 09:19:45

Input Set : A:\5470-451.ST25.TXT

Output Set : N:\CRF4\07202006\I611949.raw

```

312 <221> NAME/KEY: CDS
313 <222> LOCATION: (692)..(2197)
315 <400> SEQUENCE: 4
316 gaggacgcca tgattggttg gcgctggggc ggcggacggt ggaagggcct ggcgagtcta      60
318 ggtttttacgc ctgtgctgga ctttctcctt ccatgtttcc aggccgtggg gggctacaga      120
320 gggcgagaag tcggctcagc ggaaacctgg atttggttct aagccgtggg gttgagaagg      180
322 ggtgaccgga agtgatcgtg ggactgaccg gaagcgaggc ctggagggga aagagagagc      240
324 gagacctggg agggaggggg cctccagcag aaagggggcg gggaaaagggt gcaaaagcag      300
326 cgtgggagcg ccgggctggc ttcctgcggc tgctgctggt ctgactggga agcagcaagc      360
328 caccactacg aactctcaag aggagtggga gtgcgggagt ccagagctgc ctctgggaag      420
330 tctgcagtag ttgagcaaag gggctctcac gttcctgaga gctgggcagg ggggattttg      480
332 gaacctgggg cagccaagaa cgagcagcca agggtagcgg agattagtgt tgcacagagc      540
334 agtgctggtc gggcttgggg gtggctggtg ggcactgcgt gggaaacctt ggttttagt      600
336 tttcttggtt tgcgttactc ctggtgggta gaattacctt ccgcgccttt gtacaagaca      660
338 cgggtgtctcc tggggcaagg aaggagccag g atg gcc tgg gct ctg aag ctg      712
339                                     Met Ala Trp Ala Leu Lys Leu
340                                     1               5
342 cct ctg gcc gac gaa gtg att gaa tcc ggg ttg gtg cag gac ttt gat      760
343 Pro Leu Ala Asp Glu Val Ile Glu Ser Gly Leu Val Gln Asp Phe Asp
344             10             15             20
346 gct agc ctg tcc ggg atc ggc cag gaa ctg ggt gct ggt gcc tat agc      808
347 Ala Ser Leu Ser Gly Ile Gly Gln Glu Leu Gly Ala Gly Ala Tyr Ser
348       25             30             35
350 atg agt gat gtc ctt gca ttg ccc att ttt aag caa gaa gag tcg agt      856
351 Met Ser Asp Val Leu Ala Leu Pro Ile Phe Lys Gln Glu Glu Ser Ser
352 40             45             50             55
354 ttg cct cct gat aat gag aat aaa atc ctg cct ttt caa tat gtg ctt      904
355 Leu Pro Pro Asp Asn Glu Asn Lys Ile Leu Pro Phe Gln Tyr Val Leu
356             60             65             70
358 tgt gct gct acc tct cca gca gtg aaa ctc cat gat gaa acc cta acg      952
359 Cys Ala Ala Thr Ser Pro Ala Val Lys Leu His Asp Glu Thr Leu Thr
360             75             80             85
362 tat ctc aat caa gga cag tct tat gaa att cga atg cta gac aat agg      1000
363 Tyr Leu Asn Gln Gly Gln Ser Tyr Glu Ile Arg Met Leu Asp Asn Arg
364             90             95             100
366 aaa ctt gga gaa ctt cca gaa att aat ggc aaa ttg gtg aag agt ata      1048
367 Lys Leu Gly Glu Leu Pro Glu Ile Asn Gly Lys Leu Val Lys Ser Ile
368       105             110             115
370 ttc cgt gtg gtg ttc cat gac aga agg ctt cag tac act gag cat cag      1096
371 Phe Arg Val Val Phe His Asp Arg Arg Leu Gln Tyr Thr Glu His Gln
372 120             125             130             135
374 cag cta gag ggc tgg agg tgg aac cga cct gga gac aga att ctt gac      1144
375 Gln Leu Glu Gly Trp Arg Trp Asn Arg Pro Gly Asp Arg Ile Leu Asp
376             140             145             150
378 ata gat atc ccg atg tct gtg ggt ata atc gat cct agg gct aat cca      1192
379 Ile Asp Ile Pro Met Ser Val Gly Ile Ile Asp Pro Arg Ala Asn Pro
380             155             160             165
382 act caa cta aat aca gtg gag ttc ctg tgg gac cct gca aag agg aca      1240
383 Thr Gln Leu Asn Thr Val Glu Phe Leu Trp Asp Pro Ala Lys Arg Thr

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/611,949

DATE: 07/20/2006

TIME: 09:19:46

Input Set : A:\5470-451.ST25.TXT

Output Set: N:\CRF4\07202006\I611949.raw